

# Stochastic Subspace Identification: Valid Model, Asymptotics and Model Error Bounds

Quan Li, Jeffrey T. Scruggs, *Member, IEEE*,

## Abstract

This paper investigates the ability of the stochastic subspace identification technique to return a valid model from finite measurement data, its asymptotic properties as the data set becomes large, and asymptotic error bounds of the identified model (in terms of  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norms). First, a new and straightforward LMI-based approach is proposed, which returns a valid identified model even in cases where the system poles are very close to unit circle and there is insufficient data to accurately estimate the covariance matrices. The approach, which is demonstrated by numerical examples, provides an alternative to other techniques which often fail under these circumstances. Then, an explicit expression for the variance of the asymptotically normally distributed sample output covariance matrices and block-Hankel matrix are derived. From this result, together with perturbation techniques, error bounds for the state-space matrices in the innovations model are derived, for a given confidence level. This result is in turn used to derive several error bounds for the identified transfer functions, for a given confidence level. One is an explicit  $\mathcal{H}_2$  bound. Additionally, two  $\mathcal{H}_\infty$  error bounds are derived; one via perturbation analysis, and the other via an LMI-based technique.

## Index Terms

Asymptotic variance, stochastic subspace identification, positive realness,  $\mathcal{H}_2$  norm error bound,  $\mathcal{H}_\infty$  norm error bound, linear matrix inequalities.

Q. Li is currently with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: quanli@mit.edu).

J.T. Scruggs is with the Department of Civil & Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA (e-mail: jscruggs@umich.edu).

## I. INTRODUCTION

In system estimation and control design, an undoubtedly crucial issue is the identification of the stochastic part of the plant [1], [2], [3]. To obtain a nominal model for the stochastic system, stochastic subspace identification methods are often used [4], [5], [6]. However, due to model error or data insufficiency, these methods may encounter a failure mode without any valid model returned. This is especially common when the system poles are very close to the unit circle, as demonstrated experimentally in [7]. To overcome this difficulty, several approximation approaches for guaranteeing a valid model have been proposed in recent years [8], [9], [10], [11]. In [8], Mari et al. proposed an improved version of stochastic subspace identification algorithm in which linear matrix inequality (LMI) based techniques are used to constrain the identified system poles inside the unit circle as well as for multivariate covariance fitting, thus guaranteeing the solvability of the associated discrete algebraic Riccati equation (DARE) and a valid model returned. Although promising, this method relies on coprime factorization for covariance estimation. For large-dimensional multiple input multiple output (MIMO) systems, the numerical robustness of coprime factorizations may be problematic, and in any case is not well understood. Goethals et al. in [9] used regularization techniques to impose positive realness on the associated covariance model, which we will define in Section II(A). The solvability of the associated DARE and thus the feasibility of a valid model are then satisfied, although at the cost of introducing a small bias on the identified model. Different from these approaches, the present paper first exploits an equivalence between the solvability of the DARE and the nonemptiness of a convex set, and then by the Positive Real Lemma, establishes in section II(B) a more straightforward approximation approach based on LMI techniques.

To estimate how much data is required to identify a model given a model error bound (such as  $\mathcal{H}_\infty$  norm bound) with a chosen confidence level as a starting point for further robust controller design, the asymptotic analysis and perturbation methods provide fundamental tools. For identification of the deterministic part of the plant, asymptotic statistical properties for prediction error (PE) methods [12], instrumental variable (IV) methods [13], and subspace methods [14], are all well-established. However, the corresponding asymptotic properties are not thoroughly studied for the identification of the stochastic part of the plant.

Due to a finite sample size and influence of system and measurement noise, a model error

always exists between the identified and true system. To assess the quality of the identified model, a model error bound in terms of  $\mathcal{H}_2$  or  $\mathcal{H}_\infty$  norms is often useful. Moreover, robust control theory is often predicated on knowledge of an upper bound for the  $\mathcal{H}_\infty$  norm of the model error. To derive an  $\mathcal{H}_\infty$  norm model error bound for the deterministic part of a system, various identification methods are proposed by [15], [16], [17], [18], [19]. In [19], an upper bound on the  $\mathcal{H}_\infty$  norm of the model error is derived via a frequency response curve fitting procedure which minimizes a maximum amplitude criterion and guarantees the stability of the identified model. In this procedure, linear as well as nonlinear programming techniques are used. Another approach to quantify a model error, proposed in [20], establishes a framework connecting PE methods with robust control theory. In this framework, the tools of PE methods are used to quantify an uncertainty region to which robustness tools are conveniently adapted such that robustness analysis of a controller and the quality assessment of the uncertainty region are easily carried out. For the stochastic part of a system, the quantification of model error is rarely reported. One of the contributions of this paper is to derive a model error bounds in terms of both  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norms, with a confidence level given by asymptotic analysis of stochastic subspace identification.

In this paper, we first propose a new and straightforward LMI-based optimization approach to identifying the stochastic system. Then for the identified stochastic system, we investigate its asymptotic behavior using asymptotic analysis and perturbation methods. Particularly, we derive the asymptotic Frobenius norm (F-norm) error bounds of the state-space matrices in innovations model. With these asymptotic properties, we derive the explicit expressions of the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bounds for the identified system, for a given confidence level. We also propose an LMI-based approach to computing the  $\mathcal{H}_\infty$  norm model error bound. In order to verify our analytical results, we first apply them to identify an innovations model for a plant with poles are very close to unit circle, and for which the amount of measurement data is insufficient to accurately evaluate output covariances. Then, we propagate these derived bounds, to estimate  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  model error bounds for the identified system. Based on simulation results, it is shown that the improved stochastic subspace identification procedure guarantees a valid model returned, together with  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bounds with a chosen high confidence level.

The outline of this paper is as follows. In Section II, the stochastic subspace identification

procedure will be outlined, and the equivalence of its failure mode with the emptiness of a convex set will be illustrated, which motivates an LMI-based optimization approach for guaranteeing a valid model returned. The asymptotic analysis and perturbation method will be given in Section III, deriving the asymptotic distributions of state-space matrices in the associated covariance model, and the asymptotic F-norm error bounds of the state-space matrices in innovations model with a chosen confidence level. In Section IV, these asymptotic results are combined to derive the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bounds of the identified system with a chosen confidence level. Two typical numerical examples will be presented in Section V, and the conclusions will be drawn in Section VI. For brevity, hereafter we use  $I$  or  $I_n \in \mathcal{R}^{n \times n}$  to denote the identity matrix with compatible dimension,  $\tilde{(\cdot)}$  to denote the estimated value of  $(\cdot)$  from the measurement data,  $\delta(\cdot) = \tilde{(\cdot)} - (\cdot)$  to denote the perturbation of  $(\cdot)$  due to finite data samples,  $o(\cdot)$  to denote the higher order perturbation satisfying  $o(\cdot)/(\cdot) \rightarrow 0$  as  $(\cdot) \rightarrow 0$ , and  $\dot{=}$  to denote the first-order approximation due to the perturbations from the finite sample.  $\text{vec}(\cdot)$ ,  $\otimes$ , and  $\text{tr}(\cdot)$  represent the columnwise vectorization of a matrix, Kronecker product and the trace of a matrix, respectively.  $(\cdot)^*$ ,  $(\cdot)^T$  and  $(\cdot)^H$  represent conjugate, transpose and conjugate transpose, respectively.  $\|\cdot\|_F$  and  $\|\cdot\|_2$  represent F-norm and 2-norm, respectively.

## II. AN IMPROVED STOCHASTIC SUBSPACE IDENTIFICATION ALGORITHM

In this section, we will first recall the standard stochastic identification [4], [8] in which the associated DARE may be unsolvable especially when measurement data is insufficient to estimate stationary output covariances, and the system poles are very closed to the unit circle. To overcome this difficulty, we establish an equivalence between the solvability of the DARE and the positive realness of an associated covariance model, and then propose an LMI-based approach to impose positive realness on the covariance model, thus guaranteeing a valid model is returned.

### A. Standard subspace identification procedure

Assume the following state-space innovations model is a minimal realization of the vector stochastic process  $\{y_k\}$

$$\begin{aligned} x_{k+1} &= Ax_k + KQ^{\frac{1}{2}}e_k \\ y_k &= Cx_k + Q^{\frac{1}{2}}e_k \end{aligned} \tag{1}$$

where  $x_k \in \mathcal{R}^{n_x \times 1}$ ;  $y_k \in \mathcal{R}^{n_y \times 1}$ ;  $K$  is the Kalman gain;  $Q \in \mathcal{R}^{n_y \times n_y}$  is the innovations covariance matrix;  $e_k \in \mathcal{R}^{n_y \times 1}$  are the normalized white innovations with covariance matrix  $E[e_k e_k^T]$  equal to  $I_{n_y}$ . We assume zero-mean processes throughout. For  $k \in \mathcal{Z}_{\geq 0}$ , denote the process covariances as

$$R_k = E[y(t)y^T(t-k)] \quad (2)$$

According to the minimal realization assumption, the following observability matrix  $\Omega$  and controllability matrix  $\Gamma$  are both in full rank.

$$\Omega = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{m-1} \end{bmatrix}, \Gamma = [D \quad AD \quad \cdots \quad A^{m-1}D] \quad (3)$$

where  $D = E[x_{k+1}y_k^T]$  and  $m > n_x$ . Noting that the sequence  $\{R_i\}$  are the Markov parameters of the system  $\{A, D, C, R_0\}$ , the following factorization of the block-Hankel matrix of the output covariance holds:

$$H = \begin{bmatrix} R_1 & R_2 & \cdots & R_m \\ R_2 & R_3 & \cdots & R_{m+1} \\ \vdots & & \ddots & \vdots \\ R_m & R_{m+1} & \cdots & R_{2m-1} \end{bmatrix} = \Omega\Gamma \quad (4)$$

From (3) and (4), the rank properties of  $\Omega$  and  $\Gamma$  imply that

$$\text{rank}(H) = n_x, \quad \text{for } m > n_x \quad (5)$$

We refer to  $(A, D, C, R_0/2)$  as the covariance model associated with the innovations model (1).

From singular value decomposition (SVD) of  $H$ , we have

$$\begin{aligned} \Omega\Gamma &= [U_s \quad U_n] \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} [V_s \quad V_n]^T \\ &= U_s \Lambda_s V_s^T \end{aligned} \quad (6)$$

where the diagonal matrix  $\Lambda_s \in \mathcal{R}^{n_x \times n_x}$  is nonsingular. (6) indicates the realizations of  $\Omega$  and  $\Gamma$  as

$$\Omega = U_s \Lambda_s^{\frac{1}{2}} T, \quad \Gamma = T^{-1} \Lambda_s^{\frac{1}{2}} V_s^T \quad (7)$$

where  $T \in \mathcal{R}^{n_x \times n_x}$  is any (nonsingular) similarity transformation matrix. Hereafter we set  $T = I$ . In view of (3), it is indicated that  $C$  is the first  $n_y$  rows of  $\Omega$ , and  $D$  is the first  $n_y$  columns of  $\Gamma$ .  $A$  can then be solved by the following overdetermined linear equations:

$$\overline{\Omega}A = \underline{\Omega} \quad (8)$$

where  $\overline{\Omega}$  and  $\underline{\Omega}$  are respectively the first  $m-1$  matrix blocks and the last  $m-1$  matrix blocks of  $\Omega$ . We then solve the discrete algebraic Ricatti equation (DARE) to obtain a positive definite covariance matrix  $P = E[x_k x_k^T]$

$$P = APA^T + (D - APC^T)(R_0 - CPC^T)^{-1}(D - APC^T)^T \quad (9)$$

and  $K$  and  $Q$  are then given by

$$\begin{aligned} Q &= R_0 - CPC^T \\ K &= (D - APC^T)Q^{-1} \end{aligned} \quad (10)$$

### B. Correcting for errors due to finite measurement data

In practice,  $R_k$  is estimated by the empirical (i.e., statistical) sample output covariance

$$\tilde{R}_k = \frac{1}{N-k} \sum_{t=k+1}^N y_t y_{t-k}^T, \quad k = 0, 1, \dots, 2m-1 \quad (11)$$

The above algorithm is then implemented assuming  $R_k \approx \tilde{R}_k$ . The resultant estimations for  $\{A, C, Q, K\}$  are denoted  $\{\tilde{A}, \tilde{C}, \tilde{Q}, \tilde{K}\}$ . As  $N \rightarrow \infty$ , the estimations converge to the true values, but for  $N$  finite, the inevitable existence of estimation error in  $R_k$  leads to two problems.

The first potential problem is that  $\tilde{A}$ , as determined as the least-squares solution to (8), may be unstable. If this is the case, a simple LMI-based correction proposed in [8] can be used to recover stability:

$$\begin{aligned} \min_{\hat{A}, W} \quad & \|(\hat{A} - \tilde{A})W\|_F \\ \text{s.t.} \quad & W - \hat{A}W\hat{A}^T > 0 \\ & W > 0 \end{aligned} \quad (12)$$

where  $\hat{A}$  is the adjusted (and asymptotically stable) approximation of  $\tilde{A}$ .

The second potential problem is that the DARE (9) may fail to yield a positive-definite, real solution  $P$ , thus prohibiting the derivation of approximate innovations model parameters  $\{\tilde{Q}, \tilde{K}\}$

as in (10). Failures are especially likely for large dimensional systems with poles close to the unit circle. The following result reveals the equivalence between the solvability of DARE and the positive realness of the covariance model. For the theorem, we assume a covariance model  $\{\hat{A}, \tilde{D}, \tilde{C}, \tilde{R}_0/2\}$  has been derived via the procedure above.

*Proposition 1:* Given a symmetric, positive definite matrix  $\tilde{R}_0$  and a controllable and observable covariance model  $\{\hat{A}, \tilde{D}, \tilde{C}, \tilde{R}_0/2\}$ , the associated Riccati equation (9) has a positive definite solution  $P$  with innovations covariance  $\tilde{Q} = \tilde{R}_0 - \tilde{C}P\tilde{C}^T \geq 0$  if and only if the set  $\mathcal{P}$ , defined as

$$\mathcal{P} = \left\{ P \mid \begin{bmatrix} P - \hat{A}P\hat{A}^T & \tilde{D} - \hat{A}P\tilde{C}^T \\ \tilde{D}^T - \tilde{C}P\hat{A}^T & \tilde{R}_0 - \tilde{C}P\tilde{C}^T \end{bmatrix} \geq 0, P > 0 \right\} \quad (13)$$

is nonempty.

*Proof:* [4] and [10] have shown that given a controllable and observable model, (9) has a positive definite solution  $P$  with  $\tilde{R}_0 - \tilde{C}P\tilde{C}^T \geq 0$  if and only if  $\tilde{C}(zI - \hat{A})^{-1}\tilde{D} + \tilde{R}_0/2$  is positive real. Thus, we have  $\tilde{D}^T(zI - \hat{A}^T)^{-1}\tilde{C}^T + \tilde{R}_0/2$  is also positive real. Applying the Positive Real Lemma obtains the conclusion. ■

Thus, the failure of DARE indicates a null set  $\mathcal{P}$  resulting from the estimate  $(\hat{A}, \tilde{D}, \tilde{C}, \tilde{R}_0)$ . This suggests that a feasible member in  $\mathcal{P}$  be approximated by solving the following optimization problem

$$\begin{aligned} \left( P, \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{bmatrix} \right) = \arg \min_{P, \Phi_{11}, \Phi_{12}, \Phi_{22}} & \left\| \begin{bmatrix} P - \hat{A}P\hat{A}^T & \tilde{D} - \hat{A}P\tilde{C}^T \\ \tilde{D}^T - \tilde{C}P\hat{A}^T & \tilde{R}_0 - \tilde{C}P\tilde{C}^T \end{bmatrix} - \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{bmatrix} \right\| \\ \text{s.t.} & \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{bmatrix} \geq 0 \\ & P > 0 \end{aligned} \quad (14)$$

in which we can use 2-norm or F-norm to express it as a semidefinite program. For simplicity,

choosing 2-norm in (14) gives the following LMI problem

$$\min \quad \lambda$$

$$\begin{bmatrix} \lambda \mathbf{I} & \mathbf{0} & P - \hat{A}P\hat{A}^T - \Phi_{11} & \tilde{D} - \hat{A}P\tilde{C}^T - \Phi_{12} \\ & \lambda \mathbf{I} & \tilde{D}^T - \tilde{C}P\hat{A}^T - \Phi_{12}^T & \tilde{R}_0 - \tilde{C}P\tilde{C}^T - \Phi_{22} \\ & & \mathbf{I} & \mathbf{0} \\ (sym) & & & \mathbf{I} \end{bmatrix} > 0 \quad (15)$$

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{bmatrix} \geq 0 \quad (16)$$

$$P > 0 \quad (17)$$

where the variables are the positive definite matrix  $P$ , the nonnegative definite matrices  $\Phi_{11}$  and  $\Phi_{22}$ , the non-symmetric matrix  $\Phi_{12}$ , and the scalar  $\lambda$ .

For convenience, we keep the estimated  $\hat{A}$  and  $\tilde{C}$  unchanged, and adjust  $\tilde{D}$  and  $\tilde{R}_0$  for guaranteeing that a positive definite solution to (9) exists. After solving (14), we obtain  $\bar{P}$  as the solution to the following discrete Lyapunov equation:

$$\bar{P} = \hat{A}\bar{P}\hat{A}^T + \Phi_{11} \quad (18)$$

and then the new  $\hat{D}$  and  $\hat{R}_0$  are adjusted from  $\tilde{D}$  and  $\tilde{R}_0$  by

$$\hat{D} = \hat{A}\bar{P}\tilde{C}^T + \Phi_{12} \quad (19)$$

$$\hat{R}_0 = \tilde{C}\bar{P}\tilde{C}^T + \Phi_{22} \quad (20)$$

Substituting the adjusted  $\hat{D}$  and  $\hat{R}_0$  to DARE (9), it is guaranteed that DARE has a positive definite solution  $\hat{P}$ , and finally, the estimated  $\hat{K}$  and  $\hat{Q}$  are updated by

$$\hat{Q} = \hat{R}_0 - \tilde{C}\hat{P}\tilde{C}^T \quad (21)$$

$$\hat{K} = (\hat{D} - \hat{A}\hat{P}\tilde{C}^T)\hat{Q}^{-1} \quad (22)$$

We thus arrive at adjusted parameters  $\{\hat{A}, \hat{K}, \tilde{C}, \hat{Q}\}$  as the approximation to the innovations model parameters from (1).



### III. ASYMPTOTIC AND PERTURBATION ANALYSIS OF STOCHASTIC SUBSPACE IDENTIFICATION

In this section, we will derive the asymptotic distributions of the empirical sample output covariance matrices, with its variance expressed in terms of the power spectral density (PSD) of the vector stochastic process  $y_k$ . Later we will use matrix perturbation analysis to derive the asymptotic variances of the asymptotically normally distributed state-space matrices in the covariance model, and the perturbed state-space matrices in the innovations model in terms of the perturbations  $\delta R_0$  and  $\delta H$ . Using Chi-square cumulative distribution function, the F-norm error bounds of state-space matrices in the innovations model are given with a chosen confidence level.

#### A. Asymptotic Distributions of the Sample Output Covariance Matrices

To facilitate the deduction of asymptotic distributions of the empirical sample output covariance estimates of the vector stochastic process  $\{y_k\}$ , we introduce a Lemma which illustrates that the expectation of the mixed product of scalar and vector random variables can be expressed in terms of first- and second-order moments. This Lemma, as a special case of theorem 1 in [21], gives a useful tool to calculate the fourth order moment.

*Lemma 1:* If  $x_1, x_2 \in \mathcal{R}$  and  $X_3, X_4 \in \mathcal{R}^{n \times 1}$  have jointly gaussian distributions, then

$$\begin{aligned} E[x_1 x_2 X_3 X_4^T] &= E[x_1 x_2] E[X_3 X_4^T] + E[x_1 X_3] E[x_2 X_4^T] + E[x_2 X_3] E[x_1 X_4^T] \\ &\quad - 2E[x_1]E[x_2]E[X_3]E[X_4^T] \end{aligned} \quad (23)$$

We can now propose a theorem which reveals the asymptotic distributions of the empirical sample output covariance estimates of the vector stochastic process  $\{y_k\}$ .

*Theorem 1:* Consider the empirical sample covariance  $\tilde{R}_0$  for the vector stochastic process  $\{y_k\}$  given by (1) and the empirical Hankel matrix  $\tilde{H}$  given by substituting each  $R_k$  in (4) by  $\tilde{R}_k$  in (11). Then  $\tilde{R}_0$  and  $\tilde{H}$  are both asymptotically normally distributed

$$\sqrt{N} \left( \begin{bmatrix} \text{vec}(\tilde{R}_0 - R_0) \\ \text{vec}(\tilde{H} - H) \end{bmatrix} \right) \longrightarrow \mathcal{N}(0, \mathcal{P}_{R_0, H}) \quad (24)$$

where the variance matrix  $\mathcal{P}_{R_0,H} = \begin{bmatrix} P_{R_0} & P_{R_0H} \\ P_{R_0H}^T & P_H \end{bmatrix}$  with the block matrices  $P_{R_0}$ ,  $P_{R_0H}$  and  $P_H$  given by

$$P_{R_0} = (I_{n_x^2} + K_{n_y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} S_y^*(\omega) \otimes S_y(\omega) d\omega \quad (25)$$

$$P_{R_0H} = (I_{n_x^2} + K_{n_y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} (E_2^H \otimes S_y(\omega))^* \otimes (E_1^H \otimes S_y(\omega)) d\omega \quad (26)$$

$$P_H = (I_{n_x^2 m^2} + K_{n_y m}) \frac{1}{2\pi} \int_{-\pi}^{\pi} ((E_2 E_2^H) \otimes S_y(\omega))^* \otimes ((E_1 E_1^H) \otimes S_y(\omega)) d\omega \quad (27)$$

where  $S_y(\omega)$  is the PSD of the vector stochastic process  $\{y_k\}$ ;  $K_n$  is a permutation matrix satisfying

$$K_n(\mathbf{e}_i \otimes \mathbf{e}_j) = \mathbf{e}_j \otimes \mathbf{e}_i \quad (28)$$

;  $E_1 = [1 \ e^{i\omega} \ \dots \ e^{i\omega(m-1)}]^T$ ;  $E_2 = [e^{-i\omega} \ e^{-2i\omega} \ \dots \ e^{-i\omega m}]^T$ ;  $\mathbf{e}_i$  is a  $n$ -dimensional unit vector with the  $i$ th element be 1 and others 0.

*Proof:* We first proof (25), and then extend the proof to obtain (26) and (27). We have the covariance of  $\text{vec}(\tilde{R}_0 - R_0)$

$$\begin{aligned} & \text{Cov}(\text{vec}(\tilde{R}_0 - R_0), \text{vec}(\tilde{R}_0 - R_0)) \\ &= \frac{1}{N^2} \sum_{1 \leq k, h \leq N} E \left[ (y_k \otimes y_k - \text{vec}(R_0)) (y_h \otimes y_h - \text{vec}(R_0))^T \right] \end{aligned} \quad (29)$$

where we derive the second line using the fact  $\text{vec}(y_k y_k^T) = y_k \otimes y_k$ . Each entry in the sum in (29) is

$$\begin{aligned} & E \left[ (y_k \otimes y_k - \text{vec}(R_0)) (y_h \otimes y_h - \text{vec}(R_0))^T \right] \\ &= E \left[ (y_k y_h^T) \otimes (y_k y_h^T) \right] - \text{vec}(R_0) \text{vec}(R_0)^T \\ &= \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes E \left( y_k^{(i)} y_h^{(j)} y_k y_h^T \right) - \text{vec}(R_0) \text{vec}(R_0)^T \\ &= \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes E(y_k^{(i)} y_h^{(j)}) E(y_k y_h^T) + \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes E(y_k^{(i)} y_k) E(y_h^{(j)} y_h^T) \\ &\quad + \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes E(y_h^{(j)} y_k) E(y_k^{(i)} y_h^T) - \text{vec}(R_0) \text{vec}(R_0)^T \end{aligned} \quad (30)$$

where we use Lemma 1 to derive the fourth line from the third line. For the first term in (30), we have

$$\begin{aligned}
& \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes E(y_k^{(i)} y_h^{(j)}) E(y_k y_h^T) \\
&= \sum_{i,j,m,n} (\mathbf{e}_i \mathbf{e}_j^T) \otimes (\mathbf{e}_m \mathbf{e}_n^T) E(y_k^{(i)} y_h^{(j)}) E(y_k^{(m)} y_h^{(n)}) \\
&= \sum_{i,j,m,n} (\mathbf{e}_i \otimes \mathbf{e}_m) (\mathbf{e}_j^T \otimes \mathbf{e}_n^T) E(y_k^{(i)} y_h^{(j)}) E(y_k^{(m)} y_h^{(n)}) \tag{31}
\end{aligned}$$

For the second term in (30), we have

$$\begin{aligned}
& \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes (E(y_k^{(i)} y_k) E(y_h^{(j)} y_h^T)) \\
&= \sum_{i,j} \left( \mathbf{e}_i \otimes E(y_k^{(i)} y_k) \right) \left( \mathbf{e}_j \otimes E(y_h^{(j)} y_h) \right)^T \\
&= \sum_{i,j} \text{vec}(E(y_k^{(i)} y_k) \mathbf{e}_i^T) \text{vec}^T(E(y_h^{(j)} y_h) \mathbf{e}_j^T) \\
&= \text{vec}(E(y_k y_k^T)) \text{vec}^T(E(y_h y_h^T)) \\
&= \text{vec}(R_0) \text{vec}(R_0)^T \tag{32}
\end{aligned}$$

For the third term in (30), we have

$$\begin{aligned}
& \sum_{i,j} \mathbf{e}_i \mathbf{e}_j^T \otimes E(y_h^{(j)} y_k) E(y_k^{(i)} y_h^T) \\
&= \sum_{i,j,m,n} ((\mathbf{e}_i \mathbf{e}_j^T) \otimes (\mathbf{e}_m \mathbf{e}_n^T)) E(y_k^{(m)} y_h^{(j)}) E(y_k^{(i)} y_h^{(n)}) \\
&= \sum_{i,j,m,n} ((\mathbf{e}_m \mathbf{e}_j^T) \otimes (\mathbf{e}_i \mathbf{e}_n^T)) E(y_k^{(i)} y_h^{(j)}) E(y_k^{(m)} y_h^{(n)}) \\
&= \sum_{i,j,m,n} (\mathbf{e}_m \otimes \mathbf{e}_i) (\mathbf{e}_j^T \otimes \mathbf{e}_n^T) E(y_k^{(i)} y_h^{(j)}) E(y_k^{(m)} y_h^{(n)}) \\
&= K_{n_y} \sum_{i,j,m,n} (\mathbf{e}_i \otimes \mathbf{e}_m) (\mathbf{e}_j^T \otimes \mathbf{e}_n^T) E(y_k^{(i)} y_h^{(j)}) E(y_k^{(m)} y_h^{(n)}) \tag{33}
\end{aligned}$$

where we switch the indices  $i$  and  $m$  to derive the third line from the second line. Substituting (31)  $\sim$  (33) to (30), we have

$$\begin{aligned}
& E \left[ (y_k \otimes y_k - \text{vec}(R_0)) (y_h \otimes y_h - \text{vec}(R_0))^T \right] \\
&= (I_{n_x^2} + K_{n_y}) \sum_{i,j,m,n} (\mathbf{e}_i \otimes \mathbf{e}_m) (\mathbf{e}_j^T \otimes \mathbf{e}_n^T) E(y_k^{(i)} y_h^{(j)}) E(y_k^{(m)} y_h^{(n)}) \\
&= (I_{n_x^2} + K_{n_y}) E(y_k y_h^T) \otimes E(y_k y_h^T) \\
&= (I_{n_x^2} + K_{n_y}) R_{k-h} \otimes R_{k-h}
\end{aligned} \tag{34}$$

Substituting (34) to (29), we have

$$\begin{aligned}
& \text{Cov}(\text{vec}(\tilde{R}_0 - R_0), \text{vec}(\tilde{R}_0 - R_0)) \\
&= \frac{1}{N^2} (I_{n_x^2} + K_{n_y}) \sum_{1 \leq k, h \leq N} R_{k-h} \otimes R_{k-h} \\
&= \frac{1}{N} (I_{n_x^2} + K_{n_y}) \sum_{\tau} R(\tau) \otimes R(\tau) \\
&= \frac{1}{N} (I_{n_x^2} + K_{n_y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{\tau_1, \tau_2} R(\tau_1) \otimes R(\tau_2) e^{-i\omega(\tau_2 - \tau_1)} d\omega \\
&= \frac{1}{N} (I_{n_x^2} + K_{n_y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{\tau_1} R(\tau_1) e^{i\omega\tau_1} \otimes \sum_{\tau_2} R(\tau_2) e^{-i\omega\tau_2} d\omega \\
&= \frac{1}{N} (I_{n_x^2} + K_{n_y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} S_y^*(\omega) \otimes S_y(\omega) d\omega
\end{aligned}$$

where  $S_y(\omega)$  denotes the PSD of the vector stochastic process  $\{y_k\}$ . Thus, (25) is concluded.

For  $P_H$ , following the same procedure we have

$$\begin{aligned}
& \text{Cov}(\text{vec}(\tilde{H} - H), \text{vec}(\tilde{H} - H)) \\
&= \text{Cov} \left( \text{vec} \left( \frac{1}{N} \sum_{k=1}^N V_{1,k} V_{2,k}^T - H \right), \text{vec} \left( \frac{1}{N} \sum_{k=1}^N V_{1,k} V_{2,k}^T - H \right) \right) \\
&= \frac{1}{N} (I_{n_x^2 m^2} + K_{n_y m}) \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{V_2}^*(\omega) \otimes S_{V_1}(\omega) d\omega
\end{aligned} \tag{35}$$

where  $V_{1,k} = [y_k^T \ y_{k+1}^T \ \dots \ y_{k+m-1}^T]^T$  and  $V_{2,k} = [y_{k-1}^T \ y_{k-2}^T \ \dots \ y_{k-m}^T]^T$ ;  $S_{V_1}(\omega)$  and  $S_{V_2}(\omega)$  denotes the spectral densities of  $V_{1,k}$  and  $V_{2,k}$  respectively, and can be obtained by

$$\begin{aligned}
S_{V_1}(\omega) &= (E_1 \otimes I_{n_x}) S_y(\omega) (E_1^H \otimes I_{n_x}) \\
&= (E_1 E_1^H) \otimes S_y(\omega)
\end{aligned} \tag{36}$$

Likewise, we have

$$\begin{aligned} S_{V_2}(\omega) &= (E_2 \otimes I_{n_x}) S_y(\omega) (E_2^H \otimes I_{n_x}) \\ &= (E_2 E_2^H) \otimes S_y(\omega) \end{aligned} \quad (37)$$

Substituting (36) and (37) to (35) gives (27). For  $P_{R_0 H}$ , following the same procedure we have that

$$\begin{aligned} P_{R_0 H} &= \text{Cov} \left( \text{vec}(\tilde{R}_0 - R_0), \text{vec}(\tilde{H} - H) \right) \\ &= \text{Cov} \left( \text{vec} \left( \frac{1}{N} \sum_{k=1}^N y_k y_k^T - R_0 \right), \text{vec} \left( \frac{1}{N} \sum_{k=1}^N V_{1,k} V_{2,k}^T - H \right) \right) \\ &= \frac{1}{N} (I_{n_x^2} + K_{n_y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{yV_2}^*(\omega) \otimes S_{yV_1}(\omega) d\omega \end{aligned} \quad (38)$$

where  $S_{yV_2} = E_2^H \otimes S_y(\omega)$  is the cross spectral density of  $\{y_k\}$  and  $\{V_{2,k}\}$ , and  $S_{yV_1} = E_1^H \otimes S_y(\omega)$  is the cross spectral density of  $\{y_k\}$  and  $\{V_{1,k}\}$ . Likewise, we have (26). Proposition 7.3.2 ~ 7.3.4 in [22], as well as theorem 1 in [23], claim that the sample covariances  $\tilde{R}_k$  in (11) are normally distributed as the measurement data size  $N$  goes to infinity. Thus, (24) is concluded. ■

Remark 1: the PSD  $S_y(\omega)$  of vector stochastic process  $\{y_k\}$  can be computed by the transfer function  $G_e(z)$  from the normalized innovations to output, i.e.

$$S_y(\omega) = G_e(e^{i\omega}) G_e^T(e^{-i\omega}) \quad (39)$$

where

$$G_e(z) = C(zI - A)^{-1} K Q^{\frac{1}{2}} + Q^{\frac{1}{2}} \quad (40)$$

### B. Asymptotic Distributions of the State-space Matrices in the Covariance Model

Based on the asymptotic distributions of the empirical sample covariances of the vector stochastic process  $\{y_k\}$ , perturbation analysis of SVD is firstly applied to derive the asymptotic distributions of the controllability matrix  $\Gamma$  and observability matrix  $\Omega$ , from which we then derive the asymptotic distributions of the state space matrices in the covariance model.

Assume the true covariance matrix  $H$  has SVD

$$H = U_s \Lambda_s V_s^T + U_n \Lambda_n V_n^T \quad (41)$$

where the diagonal matrix  $\Lambda_s \in \mathcal{R}^{n_x \times n_x}$  is defined in (6) and  $\Lambda_n = \mathbf{0}$ . Due to finite data samples, there exists a perturbation  $\delta H$  in  $\tilde{H}$

$$\tilde{H} = H + \delta H \quad (42)$$

such that there exist the corresponding perturbations in the subspaces and singular values. The SVD on  $\tilde{H}$  gives

$$\tilde{H} = \tilde{U}_s \tilde{\Lambda}_s \tilde{V}_s^T + \tilde{U}_n \tilde{\Lambda}_n \tilde{V}_n^T \quad (43)$$

Due to  $\delta H$ , all the terms on the right hand side of (43) may differ from those on the right hand side of (41). [24] developed the perturbation analysis of SVD to the second order. We will apply its main theorem to analyze how  $\delta H$  influences the SVD of  $\tilde{H}$ . Accordingly, we have that

$$\begin{aligned} \tilde{\Lambda}_s &\doteq \Lambda_s + U_s^T \delta H V_s \\ \tilde{U}_s &\doteq U_s + U_n U_n^T \delta H V_s \Lambda_s^{-1} \\ \tilde{V}_s &\doteq V_s + V_n V_n^T \delta H^T U_s \Lambda_s^{-1} \end{aligned} \quad (44)$$

Also, matrix perturbation analysis obtains the first order approximation of  $\tilde{\Lambda}_s^{\frac{1}{2}}$

$$\tilde{\Lambda}_s^{\frac{1}{2}} \doteq \Lambda_s^{\frac{1}{2}} + \delta \Lambda_{sq} \quad (45)$$

where

$$\text{vec}(\delta \Lambda_{sq}) = \left( I_{n_y} \otimes \Lambda_s^{\frac{1}{2}} + \Lambda_s^{\frac{1}{2}} \otimes I_{n_y} \right)^{-1} (V_s^T \otimes U_s^T) \text{vec}(\delta H) \quad (46)$$

Algebraic manipulation yields the perturbations of the controllability and observability matrices in the covariance model in terms of  $\delta H$

$$\text{vec} \left( \tilde{U}_s \tilde{\Lambda}_s^{\frac{1}{2}} - U_s \Lambda_s^{\frac{1}{2}} \right) \doteq \Pi_1 \text{vec}(\delta H) \quad (47)$$

$$\text{vec} \left( \tilde{\Lambda}_s^{\frac{1}{2}} \tilde{V}_s^T - \Lambda_s^{\frac{1}{2}} V_s^T \right) \doteq \Pi_2 \text{vec}(\delta H) \quad (48)$$

where

$$\begin{aligned} \Pi_1 &= \left[ (I_{n_y} \otimes U_s) \left( I_{n_y} \otimes \Lambda_s^{\frac{1}{2}} + \Lambda_s^{\frac{1}{2}} \otimes I_{n_y} \right)^{-1} (V_s^T \otimes U_s^T) + \left( \Lambda_s^{-\frac{1}{2}} V_s^T \right) \otimes (U_n U_n^T) \right] \\ \Pi_2 &= \left[ (V_s \otimes I_{n_y}) \left( I_{n_y} \otimes \Lambda_s^{\frac{1}{2}} + \Lambda_s^{\frac{1}{2}} \otimes I_{n_y} \right)^{-1} (V_s^T \otimes U_s^T) + (V_n V_n^T) \otimes \left( \Lambda_s^{-\frac{1}{2}} U_s^T \right) \right] \end{aligned} \quad (49)$$

From theorem 1 and (47) ~ (48), we have the following proposition which reveals the asymptotic distributions of the observability and controllability matrices in the covariance model.

*Proposition 2:* Consider the estimated observability matrix  $\tilde{\Omega}$  and controllability matrix  $\tilde{\Gamma}$  given by (3) for  $T = I_{n_x}$ . They are asymptotically normally distributed

$$\begin{aligned}\sqrt{N}\text{vec}(\tilde{\Omega} - \Omega) &\longrightarrow \mathcal{N}(0, P_{\Omega}) \\ \sqrt{N}\text{vec}(\tilde{\Gamma} - \Gamma) &\longrightarrow \mathcal{N}(0, P_{\Gamma})\end{aligned}\quad (50)$$

The asymptotic variance matrices are

$$P_{\Omega} = \Pi_1 P_H \Pi_1^T \quad (51)$$

$$P_{\Gamma} = \Pi_2 P_H \Pi_2^T \quad (52)$$

where  $\Pi_1$  and  $\Pi_2$  are given by (49).

By this proposition, we have the following theorem for the asymptotical distributions of the state-space matrices in the covariance model.

*Theorem 2:* Assume  $m > n$ . The state-space matrices  $\tilde{A}$ ,  $\tilde{C}$ , and  $\tilde{D}$  estimated from Section II(A) are asymptotically normally distributed

$$\begin{aligned}\sqrt{N}\text{vec}(\tilde{A} - A) &\longrightarrow \mathcal{N}(0, P_A) \\ \sqrt{N}\text{vec}(\tilde{C} - C) &\longrightarrow \mathcal{N}(0, P_C) \\ \sqrt{N}\text{vec}(\tilde{D} - D) &\longrightarrow \mathcal{N}(0, P_D)\end{aligned}\quad (53)$$

where

$$\begin{aligned}P_A &= \Xi P_{\Omega} \Xi^T, \quad P_C = (I_{n_x} \otimes \Phi_3) P_{\Omega} (I_{n_x} \otimes \Phi_3)^T \\ P_D &= (\Phi_4^T \otimes I_{n_x}) P_{\Gamma} (\Phi_4^T \otimes I_{n_x})^T\end{aligned}\quad (54)$$

and

$$\begin{aligned}\Phi_1 &= [I, 0] \in \mathcal{R}^{(m-1)n_y \times mn_y}, \quad \Phi_2 = [0, I] \in \mathcal{R}^{(m-1)n_y \times mn_y} \\ \Phi_3 &= [I, 0] \in \mathcal{R}^{n_y \times mn_y}, \quad \Phi_4 = [I, 0]^T \in \mathcal{R}^{mn_y \times n_y} \\ \Xi &= I_{n_x} \otimes \left( (\Omega^T \Phi_1^T \Phi_1 \Omega)^{-1} \Omega^T \Phi_1^T \Phi_2 \right) - A^T \otimes \left( (\Omega^T \Phi_1^T \Phi_1 \Omega)^{-1} \Omega^T \Phi_1^T \Phi_1 \right)\end{aligned}\quad (55)$$

*Proof:* From the structure of (3), we have

$$\begin{aligned}\text{vec}(\tilde{C} - C) &= \text{vec}(\Phi_3 \delta \Omega) = (I_{n_x} \otimes \Phi_3) \text{vec}(\delta \Omega) \\ \text{vec}(\tilde{D} - D) &= \text{vec}(\delta \Gamma \Phi_4) = (\Phi_4^T \otimes I_{n_x}) \text{vec}(\delta \Gamma)\end{aligned}\quad (56)$$

From (8), the least-square solution of the overdetermined linear equation for  $\tilde{A}$  is

$$\tilde{A} = \left[ (\Phi_1 \tilde{\Omega})^T \Phi_1 \tilde{\Omega} \right]^{-1} (\Phi_1 \tilde{\Omega})^T \Phi_2 \tilde{\Omega} \quad (57)$$

First-order approximation of the right-hand side in (57) and then vectorizing  $\tilde{A} - A$  yields

$$\text{vec}(\tilde{A} - A) \doteq \Xi \text{vec} \delta \Omega \quad (58)$$

From (56) and (58), applying (50) yields (53). ■

### C. Perturbation Analysis of the State-space Matrices in the Innovations Model

To facilitate the perturbation analysis, the general DARE (9) is transformed to its equivalent form.

$$P - A_s^T P (I - SP)^{-1} A_s - M = 0 \quad (59)$$

where  $M = DR_0^{-1} D^T$ ,  $A_s = A^T - C^T R_0^{-1} D^T$  and  $S = C^T R_0^{-1} C$ . In terms of  $\delta A_s$ ,  $\delta M$  and  $\delta S$ , [25] gives the first-order perturbation  $\delta P$

$$\text{vec}(\delta P) \doteq J_1^{-1} \text{vec}(\delta M) + J_1^{-1} J_2 \text{vec}(\delta A_s) + J_1^{-1} J_3 \text{vec}(\delta S) \quad (60)$$

where  $J_1 = I_{n_x^2} - A_c^T \otimes A_c^T$ ,  $J_2 = (A_c^T P \otimes I_{n_x}) K_{n_x} + I_{n_x} \otimes A_c^T P$ ,  $J_3 = A_c^T P \otimes A_c^T P$  and  $A_c = (I - SP)^{-1} A_s$ . Also from the expressions of  $A_s$ ,  $M$  and  $S$ , we have that

$$\begin{aligned} \text{vec}(\delta M) &\doteq J_4 \text{vec}(\delta R_0) + J_5 \text{vec}(\delta D) \\ \text{vec}(\delta A_s) &\doteq K_{n_x} \text{vec}(\delta A) + J_6 \text{vec}(\delta C) + J_7 \text{vec}(\delta R_0) + J_8 \text{vec}(\delta D) \\ \text{vec}(\delta S) &\doteq J_9 \text{vec}(\delta R_0) + J_{10} \text{vec}(\delta C) \end{aligned} \quad (61)$$

where

$$\begin{aligned} J_4 &= -(DR_0^{-1} \otimes DR_0^{-1}), \quad J_5 = (I_{n_x} \otimes DR_0^{-1}) K_{n_x, n_y} + DR_0^{-1} \otimes I_{n_x} \\ J_6 &= -(DR_0^{-1} \otimes I_{n_x}) K_{n_y, n_x}, \quad J_7 = DR_0^{-1} \otimes C^T R_0^{-1}, \quad J_8 = -(I_{n_x} \otimes C^T R_0^{-1}) K_{n_x, n_y} \\ J_9 &= -(C^T R_0^{-1} \otimes C^T R_0^{-1}), \quad J_{10} = (C^T R_0^{-1} \otimes I_{n_x}) K_{n_y, n_x} + I_{n_x} \otimes C^T R_0^{-1} \end{aligned} \quad (62)$$

Substituting (61) to (60) obtains

$$\begin{aligned} \text{vec}(\delta P) &\doteq J_1^{-1} (J_4 + J_2 J_7 + J_3 J_9) \text{vec}(\delta R_0) + J_1^{-1} J_2 K_{n_x} \text{vec}(\delta A) + J_1^{-1} (J_2 J_6 + J_3 J_{10}) \text{vec}(\delta C) \\ &\quad + J_1^{-1} (J_5 + J_2 J_8) \text{vec}(\delta D) \end{aligned} \quad (63)$$



where from (47), (48), (56) and (58), we have that

$$\begin{aligned} \text{vec}(\delta A) &\doteq \Xi \Pi_1 [0_{m^2 n_y^2, n_y^2} \ I_{m^2 n_y^2}] \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \\ \text{vec}(\delta C) &\doteq (I_{n_x} \otimes \Phi_3) \Pi_1 [0_{m^2 n_y^2, n_y^2} \ I_{m^2 n_y^2}] \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \\ \text{vec}(\delta D) &\doteq (\Phi_4^T \otimes I_{n_x}) \Pi_2 [0_{m^2 n_y^2, n_y^2} \ I_{m^2 n_y^2}] \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \end{aligned} \quad (64)$$

Then, using (64) we express (63) as

$$\text{vec}(\delta P) \doteq \mathcal{G}_1 \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \quad (65)$$

where

$$\begin{aligned} \mathcal{G}_1 &= J_1^{-1} (J_4 + J_2 J_7 + J_3 J_9) [I_{n_y^2} \ 0_{n_y^2, m^2 n_y^2}] + J_1^{-1} J_2 K_{n_x} \Xi \Pi_1 [0_{m^2 n_y^2, n_y^2} \ I_{m^2 n_y^2}] \\ &\quad + J_1^{-1} (J_2 J_6 + J_3 J_{10}) (I_{n_x} \otimes \Phi_3) \Pi_1 [0_{m^2 n_y^2, n_y^2} \ I_{m^2 n_y^2}] \\ &\quad + J_1^{-1} (J_5 + J_2 J_8) (\Phi_4^T \otimes I_{n_x}) \Pi_2 [0_{m^2 n_y^2, n_y^2} \ I_{m^2 n_y^2}] \end{aligned} \quad (66)$$

Denote  $B = KQ^{\frac{1}{2}}$  and  $F = Q^{\frac{1}{2}}$ . Next, we will derive the explicit expressions of the perturbations  $\delta B$  and  $\delta F$ . From (10), applying perturbation analysis obtains

$$\begin{aligned} \text{vec}(\delta F) &\doteq (I \otimes Q^{\frac{1}{2}} + Q^{\frac{1}{2}} \otimes I)^{-1} [\text{vec}(\delta R_0) - (CP \otimes I + (I \otimes CP) K_{n_x, n_y}) \text{vec}(\delta C) - (C \otimes C) \text{vec}(\delta P)] \\ \text{vec}(\delta B) &\doteq (Q^{-\frac{1}{2}} \otimes I) \text{vec}(\delta D) - (Q^{-\frac{1}{2}} CP \otimes I) \text{vec}(\delta A) - (Q^{-\frac{1}{2}} C \otimes A) \text{vec}(\delta P) \\ &\quad - (Q^{-\frac{1}{2}} \otimes AP) K_{n_x, n_y} \text{vec}(\delta C) - (Q^{-\frac{1}{2}} \otimes B) (I \otimes Q^{\frac{1}{2}} + Q^{\frac{1}{2}} \otimes I)^{-1} \text{vec}(\delta Q) \end{aligned} \quad (67)$$

Also from  $Q = R_0 - CPC^T$ , applying perturbation analysis obtains

$$\text{vec}(\delta Q) \doteq \text{vec}(\delta R_0) - (C \otimes C) \text{vec}(\delta P) - [(I \otimes CP) K_{n_x, n_y} + (CP \otimes I)] \text{vec}(\delta C) \quad (68)$$

Substituting (64), (65) and (68) to (67) yields

$$\begin{aligned} \text{vec}(\delta F) &\doteq (\mathcal{G}_2 - \mathcal{G}_3 \mathcal{G}_1) \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \\ \text{vec}(\delta B) &\doteq (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1) \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \end{aligned} \quad (69)$$

where

$$\begin{aligned}
\mathcal{G}_2 &= (I \otimes Q^{1/2} + Q^{1/2} \otimes I)^{-1} \{ [I_{n_y}^2 \ 0_{n_y^2 \times m^2 n_y^2}] \\
&\quad - (CP \otimes I + (I \otimes CP)K_{n_x, n_y})(I_{n_x} \otimes \Phi_3)\Pi_1[0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}] \} \\
\mathcal{G}_3 &= (I \otimes Q^{1/2} + Q^{1/2} \otimes I)^{-1} (C \otimes C) \\
\mathcal{G}_4 &= (Q^{-\frac{1}{2}} \otimes I)(\Phi_4^T \otimes I_{n_x})\Pi_2[0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}] \\
&\quad - (Q^{-\frac{1}{2}} CP \otimes I)\Xi\Pi_1[0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}] \\
&\quad - (Q^{-\frac{1}{2}} \otimes AP)K_{n_x, n_y}(I_{n_x} \otimes \Phi_3)\Pi_1[0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}] \\
&\quad - (Q^{-\frac{1}{2}} \otimes B)(I \otimes Q^{1/2} + Q^{1/2} \otimes I)^{-1} [I_{n_y}^2 \ 0_{n_y^2 \times m^2 n_y^2}] \\
&\quad + (Q^{-\frac{1}{2}} \otimes B)(I \otimes Q^{1/2} + Q^{1/2} \otimes I)^{-1} \{ (I \otimes CP)K_{n_x, n_y} \\
&\quad + (CP \otimes I) \} (I_{n_x} \otimes \Phi_3)\Pi_1[0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}] \\
\mathcal{G}_5 &= (Q^{-\frac{1}{2}} \otimes B)(I \otimes Q^{1/2} + Q^{1/2} \otimes I)^{-1} (C \otimes C) - (Q^{-\frac{1}{2}} C \otimes A)
\end{aligned} \tag{70}$$

According to theorem 1, as the data size  $N$  is sufficiently large, we can approximately quantify the F-norm error bound of the state-space matrix perturbations  $\delta A$ ,  $\delta B$ ,  $\delta C$  and  $\delta F$  by Chi-square cumulative distribution function with a given confidence level. For  $\delta B$ , we have that

$$\begin{aligned}
\|\delta B\|_F^2 &= \text{vec}(\delta B)^T \text{vec}(\delta B) \\
&= \frac{1}{N} \mathbf{g}^T \mathcal{P}_{R_0, H}^{1/2} (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1)^T (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1) \mathcal{P}_{R_0, H}^{1/2} \mathbf{g} \\
&\leq \frac{\mathbf{g}^T \mathbf{g}}{N} \|\mathcal{P}_{R_0, H}^{1/2} (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1)^T (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1) \mathcal{P}_{R_0, H}^{1/2}\|_2
\end{aligned} \tag{71}$$

where  $\mathbf{g}$  is a normally distributed vector with variance equal to  $I_{n_y^2 + m^2 n_y^2}$ , i.e.  $\mathbf{g} \rightarrow \mathcal{N}(0, I_{n_y^2 + m^2 n_y^2})$ .

Thus, we have

$$\mathbf{g}^T \mathbf{g} \leq \chi_\alpha^2 \tag{72}$$

with the confidence level given by Chi-square cumulative distribution function  $\alpha(n_y^2 + m^2 n_y^2, \chi_\alpha^2)$  at each of the values in  $\chi_\alpha^2$  using  $n_y^2 + m^2 n_y^2$  degrees of freedom. Likewise, with this confidence

level, it is claimed that

$$\|\delta B\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|\mathcal{P}_{R_0, H}^{1/2} (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1)^T (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1) \mathcal{P}_{R_0, H}^{1/2}\|_2 \quad (73)$$

$$\|\delta A\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|\mathcal{P}_{R_0, H}^{1/2} (\Xi \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}])^T (\Xi \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}]) \mathcal{P}_{R_0, H}^{1/2}\|_2 \quad (74)$$

$$\|\delta C\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|\mathcal{P}_{R_0, H}^{1/2} ((I_{n_x} \otimes \Phi_3) \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}])^T ((I_{n_x} \otimes \Phi_3) \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}]) \mathcal{P}_{R_0, H}^{1/2}\|_2 \quad (75)$$

$$\|\delta F\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|\mathcal{P}_{R_0, H}^{1/2} (\mathcal{G}_2 - \mathcal{G}_3 \mathcal{G}_1)^T (\mathcal{G}_2 - \mathcal{G}_3 \mathcal{G}_1) \mathcal{P}_{R_0, H}^{1/2}\|_2 \quad (76)$$

Remark: provided that (9) fails due to insufficient data, (14), as well as (19) and (20), is required to adjust  $\tilde{D}$  and  $\tilde{R}_0$  for a valid model. In this case, we shall estimate the perturbations  $\delta D$  and  $\delta R_0$  by the following equations

$$\begin{aligned} \delta D &= (\hat{D} - \tilde{D}) + (\tilde{D} - D) \\ \delta R_0 &= (\hat{R}_0 - \tilde{R}_0) + (\tilde{R}_0 - R_0) \end{aligned} \quad (77)$$

where  $\hat{D} - \tilde{D}$  and  $\hat{R}_0 - \tilde{R}_0$  result from the LMI-based adjustments in (19)  $\sim$  (20), and  $\tilde{D} - D$  and  $\tilde{R}_0 - R_0$  are asymptotically normally distributed, as shown in (53) and (24).

#### IV. $\mathcal{H}_2$ AND $\mathcal{H}_\infty$ NORM MODEL ERROR BOUNDS

In this section, we will derive the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bound with a given confidence level. For brevity, we denote the true innovations model transfer function and the identified one as

$$G_e(\omega) = \left[ \begin{array}{c|c} A & B \\ \hline C & F \end{array} \right], \quad \tilde{G}_e(\omega) = \left[ \begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{F} \end{array} \right] \quad (78)$$

##### A. $\mathcal{H}_2$ Norm Model Error Bound for the Identified System

*Theorem 3:* Let  $(A, B, C, F)$  and  $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{F})$  be state-space representations of the original and identified systems, as shown in (78), such that

$$\begin{aligned} \delta A &= o(A), \quad \delta B = o(B) \\ \delta C &= o(C), \quad \delta F = o(F) \end{aligned} \quad (79)$$

Assume, without loss of generality, that the state matrices  $A$  and  $\tilde{A}$  are Hurwitz, and that the data size  $N$  is sufficiently large such that (24) approximately holds. Then, with the confidence

level  $\alpha(n_y^2 + m^2 n_y^2, \chi_\alpha^2)$ , the  $\mathcal{H}_2$  norm of the error system can be bounded by

$$\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_2}^2 \leq \|\delta F\|_F^2 + \|\delta B\|_F^2 \|\bar{P}\|_F + 2\|\bar{B}\|_F \|\delta B\|_F \|\delta P_1\|_F + \|\bar{B}\|_F^2 \|\delta P_2\|_F \quad (80)$$

where  $\delta P_1$ ,  $\delta P_2$ ,  $\delta A$ ,  $\delta B$ ,  $\delta C$  and  $\delta F$  are bounded by

$$\|\delta P_1\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|P_{R_0, H}^{\frac{1}{2}} M_1^T M_1 P_{R_0, H}^{\frac{1}{2}}\|_2 \quad (81)$$

$$\|\delta P_2\|_F \leq \|(\bar{A}^T \otimes \bar{A}^T - I_{4n_x^2})^{-1}\|_2 (2\|\bar{A}\|_F \|\delta P_1\|_F \|\delta A\|_F + \|\bar{P}\|_F \|\delta A\|_F^2 + \|\delta C\|_F^2) \quad (82)$$

$$\|\delta A\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|\mathcal{P}_{R_0, H}^{1/2} (\Xi \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}])^T (\Xi \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}]) \mathcal{P}_{R_0, H}^{1/2}\|_2 \quad (83)$$

$$\|\delta B\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|P_{R_0, H}^{\frac{1}{2}} (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1)^T (\mathcal{G}_4 + \mathcal{G}_5 \mathcal{G}_1) P_{R_0, H}^{\frac{1}{2}}\|_2 \quad (84)$$

$$\|\delta C\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|\mathcal{P}_{R_0, H}^{1/2} ((I_{n_x} \otimes \Phi_3) \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}])^T ((I_{n_x} \otimes \Phi_3) \Pi_1 [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}]) \mathcal{P}_{R_0, H}^{1/2}\|_2 \quad (85)$$

$$\|\delta F\|_F^2 \leq \frac{\chi_\alpha^2}{N} \|P_{R_0, H}^{\frac{1}{2}} (\mathcal{G}_2 - \mathcal{G}_3 \mathcal{G}_1)^T (\mathcal{G}_2 - \mathcal{G}_3 \mathcal{G}_1) P_{R_0, H}^{\frac{1}{2}}\|_2 \quad (86)$$

,  $\bar{P}$  is the positive definite solution of the following Lyapunov equation

$$\bar{A}^T \bar{P} \bar{A} - \bar{P} + \bar{C}^T \bar{C} = 0 \quad (87)$$

, and

$$\begin{aligned} \bar{A} &= \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} B \\ B \end{bmatrix}, \quad \bar{C} = [C \quad -C] \\ M_1 &= -[\bar{A}^T \otimes \bar{A}^T - I]^{-1} \{[(\bar{A}^T \bar{P} \otimes I) K_{2n_x} + (I \otimes \bar{A}^T \bar{P})] \left( \begin{bmatrix} 0 \\ I_{n_x} \end{bmatrix} \otimes \begin{bmatrix} 0 \\ I_{n_x} \end{bmatrix} \right) \Xi \Pi_1 \\ &\quad + [(I \otimes \bar{C}^T) + (\bar{C}^T \otimes I) K_{n_y, 2n_x}] \left( \begin{bmatrix} 0 \\ -I_{n_x} \end{bmatrix} \otimes I_{n_y} \right) (I_{n_x} \otimes \Phi_3) \Pi_1\} [0_{m^2 n_y^2 \times n_y^2} \ I_{m^2 n_y^2}] \end{aligned} \quad (88)$$

*Proof:* For brevity, we denote

$$\delta \bar{A} = \begin{bmatrix} 0 & 0 \\ 0 & \delta A \end{bmatrix}, \quad \delta \bar{B} = \begin{bmatrix} 0 \\ \delta B \end{bmatrix}, \quad \delta \bar{C} = [0 \quad -\delta C] \quad (89)$$

Consider the transfer function of the error system

$$G_e(\omega) - \tilde{G}_e(\omega) = \left[ \begin{array}{cc|c} A & \mathbf{0} & B \\ \mathbf{0} & \tilde{A} & \tilde{B} \\ \hline C & -\tilde{C} & F - \tilde{F} \end{array} \right] \quad (90)$$

The  $\mathcal{H}_2$  norm of  $G_e(\omega) - \tilde{G}_e(\omega)$  is computed by an algebraic approach

$$\begin{aligned} & \|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_2}^2 \\ &= \text{tr} \left[ (F - \tilde{F})^T (F - \tilde{F}) + \begin{pmatrix} B \\ \tilde{B} \end{pmatrix}^T \mathcal{P} \begin{pmatrix} B \\ \tilde{B} \end{pmatrix} \right] \\ &= \|\delta F\|_F^2 + \text{tr} \left[ \begin{pmatrix} B \\ \tilde{B} \end{pmatrix}^T \mathcal{P} \begin{pmatrix} B \\ \tilde{B} \end{pmatrix} \right] \end{aligned} \quad (91)$$

where  $\mathcal{P}$  is the solution to the following discrete Lyapunov equation

$$(\bar{A} + \delta \bar{A})^T \mathcal{P} (\bar{A} + \delta \bar{A}) - \mathcal{P} + (\bar{C} + \delta \bar{C})^T (\bar{C} + \delta \bar{C}) = 0 \quad (92)$$

Assume the asymptotic expansion of the solution  $\mathcal{P}$  is

$$\mathcal{P} = \bar{P} + \delta P_1 + \delta P_2 + o(\delta P_2) \quad (93)$$

where  $\delta P_1 = o(\bar{P})$  and  $\delta P_2 = o(\delta P_1)$ . Applying dominant balance to the perturbed equation (92) yields

$$\bar{A}^T \bar{P} \bar{A} - \bar{P} + \bar{C}^T \bar{C} = 0 \quad (94)$$

$$\bar{A}^T \delta P_1 \bar{A} - \delta P_1 + \delta \bar{A}^T \bar{P} \bar{A} + \bar{A}^T \bar{P} \delta \bar{A} + \bar{C}^T \delta \bar{C} + \delta \bar{C}^T \bar{C} = 0 \quad (95)$$

$$\bar{A}^T \delta P_2 \bar{A} - \delta P_2 + \delta \bar{A}^T \delta P_1 \bar{A} + \delta \bar{A}^T \bar{P} \delta \bar{A} + \bar{A}^T \delta P_1 \delta \bar{A} + \delta \bar{C}^T \delta \bar{C} = 0 \quad (96)$$

From (94), we have that

$$\bar{P} = \begin{bmatrix} X & -X \\ -X & X \end{bmatrix} \quad (97)$$

where  $X$  is the solution to the following discrete Lyapunov equation

$$A^T X A - X + C^T C = 0 \quad (98)$$

Denote  $\delta P_1$  as

$$\delta P_1 = \begin{bmatrix} \delta X_{11} & \delta X_{12} \\ \delta X_{12}^T & \delta X_{22} \end{bmatrix} \quad (99)$$

Substituting (97) and (99) to (95) yields

$$\begin{aligned} A^T \delta X_{11} A - \delta X_{11} &= 0 \\ A^T \delta X_{12} A - \delta X_{12} - A^T X \delta A - C^T \delta C &= 0 \\ A^T \delta X_{22} A - \delta X_{22} + \delta A^T X A + A^T X \delta A + C^T \delta C + \delta C^T C &= 0 \end{aligned} \quad (100)$$

From (100), we have that

$$\begin{aligned} \delta X_{11} &= 0 \\ \delta X_{22} &= -\delta X_{12} - \delta X_{12}^T \end{aligned} \quad (101)$$

Thus for the second term of (91), we have that

$$\begin{aligned} &\text{tr} [(\bar{B} + \delta \bar{B})^T (\bar{P} + \delta P_1 + \delta P_2 + o(\delta P_2)) (\bar{B} + \delta \bar{B})] \\ &= \text{tr} [\bar{B}^T \bar{P} \bar{B}] + \text{tr} [\delta \bar{B}^T \bar{P} \bar{B} + \bar{B}^T \delta P_1 \bar{B} + \bar{B}^T \bar{P} \delta \bar{B}] \\ &\quad + \text{tr} [\delta \bar{B}^T \bar{P} \delta \bar{B} + \delta \bar{B}^T \delta P_1 \bar{B} + \bar{B}^T \delta P_1 \delta \bar{B} + \bar{B}^T \delta P_2 \bar{B}] + o(\delta P_2) \end{aligned} \quad (102)$$

It is readily verified that in (102),

$$\begin{aligned} \text{tr} [\bar{B}^T \bar{P} \bar{B}] &= 0 \\ \text{tr} [\delta \bar{B}^T \bar{P} \bar{B} + \bar{B}^T \delta P_1 \bar{B} + \bar{B}^T \bar{P} \delta \bar{B}] &= 0 \\ \text{tr} [\delta \bar{B}^T \bar{P} \delta \bar{B} + \delta \bar{B}^T \delta P_1 \bar{B} + \bar{B}^T \delta P_1 \delta \bar{B} + \bar{B}^T \delta P_2 \bar{B}] \\ &\leq \|\delta B\|_F^2 \|\bar{P}\|_F + 2\|\bar{B}\|_F \|\delta B\|_F \|\delta P_1\|_F + \|\bar{B}\|_F^2 \|\delta P_2\|_F \end{aligned} \quad (103)$$

Next, from (95) and (96), we will derive the F-norm bounds of  $\delta P_1$  and  $\delta P_2$ , respectively. From (95), we have that

$$\begin{aligned} \text{vec}(\delta P_1) &= -[\bar{A}^T \otimes \bar{A}^T - I]^{-1} [(\bar{A}^T \bar{P} \otimes I) K_{2n_x} + (I \otimes \bar{A}^T \bar{P})] \text{vec}(\delta \bar{A}) \\ &\quad - [\bar{A}^T \otimes \bar{A}^T - I]^{-1} [(I_{2n_x} \otimes \bar{C}^T) + (\bar{C}^T \otimes I_{2n_x}) K_{n_y, 2n_x}] \text{vec}(\delta \bar{C}) \\ &= M_1 \begin{bmatrix} \text{vec}(\delta R_0) \\ \text{vec}(\delta H) \end{bmatrix} \end{aligned} \quad (104)$$

Thus, with the confidence level  $\alpha(n_y^2 + m^2 n_y^2, \chi_\alpha^2)$ , we have (81). Likewise, from (96), we have that

$$\text{vec}(\delta P_2) = -(\bar{A}^T \otimes \bar{A}^T - I)^{-1} \text{vec}(\delta \bar{A}^T \delta P_1 \bar{A} + \delta \bar{A}^T \bar{P} \delta \bar{A} + \bar{A}^T \delta P_1 \delta \bar{A} + \delta \bar{C}^T \delta \bar{C}) \quad (105)$$

Taking 2-norm on both sides in (105) yields (82) in the deduction of which we use the fact  $\|\text{vec}(\cdot)\|_2 = \|\cdot\|_F$ . From (102)  $\sim$  (105) and (91), we have the  $\mathcal{H}_2$  norm model error bound given by (80) where with the confidence level  $\alpha(n_y^2 + m^2 n_y^2, \chi_\alpha^2)$ , the F-norm bounds for  $\delta A$ ,  $\delta B$ ,  $\delta C$  and  $\delta F$  are given by (83)  $\sim$  (86). ■

### B. $\mathcal{H}_\infty$ Norm Model Error Bound for the Identified System

In this subsection, we will propose two approaches, based on perturbation analysis and LMI technique, respectively, to computing the  $\mathcal{H}_\infty$  norm model error bound for  $G_e(z)$  with a given confidence level.

Assume the data size  $N$  is sufficiently large such that (24) approximately holds. Thus, with a chosen confidence level  $\alpha(n_e^2 + m^2 n_e^2, \chi_\alpha^2)$ , we have the F-norm error bounds of the state space matrices in innovations model, shown in (83)  $\sim$  (86).

A straightforward way to derive the  $\mathcal{H}_\infty$ -norm bound of the error system is by perturbation analysis. Asymptotic expansion of the transfer function  $G_e(\omega)$  for the original system with respect to the parameters for the identified system  $\tilde{G}_e(\omega)$  in (90) yields

$$\begin{aligned} & G_e(\omega) - \tilde{G}_e(\omega) \\ &= (\tilde{C} - \delta C) \left( e^{i\omega} I - \tilde{A} + \delta A \right)^{-1} (\tilde{B} - \delta B) + \tilde{F} - \delta F - \tilde{C} \left( e^{i\omega} I - \tilde{A} \right)^{-1} \tilde{B} - \tilde{F} \\ &= -\delta C (e^{i\omega} I - \tilde{A})^{-1} \tilde{B} - \tilde{C} (e^{i\omega} I - \tilde{A})^{-1} \delta A (e^{i\omega} I - \tilde{A})^{-1} \tilde{B} \\ &\quad - \tilde{C} (e^{i\omega} I - \tilde{A})^{-1} \delta B - \delta F + o(\tilde{\Delta}) \end{aligned} \quad (106)$$

where  $\tilde{\Delta} = (\|\delta A\|_F, \|\delta B\|_F, \|\delta C\|_F, \|\delta F\|_F)^T$ . Taking  $\mathcal{H}_\infty$ -norm in both sides of (106) and then applying triangle inequality, submultiplicative inequality ( $\mathcal{H}_\infty$  norm is defined on a closed Banach space), and the fact  $\|\cdot\|_2 \leq \|\cdot\|_F$  yields

$$\begin{aligned} & \|G_e(\omega) - \tilde{G}_e(\omega)\|_\infty \\ &\leq \|\tilde{C} (e^{i\omega} I - \tilde{A})^{-1}\|_\infty \|\delta A\|_F \| (e^{i\omega} I - \tilde{A})^{-1} \tilde{B} \|_\infty + \|\delta F\|_F \\ &\quad + \|\delta C\|_F \| (e^{i\omega} I - \tilde{A})^{-1} \tilde{B} \|_\infty + \|\tilde{C} (e^{i\omega} I - \tilde{A})^{-1}\|_\infty \|\delta B\|_F + o(\|\tilde{\Delta}\|_F) \end{aligned} \quad (107)$$

With the identified state-space matrices  $\tilde{A}$ ,  $\tilde{B}$ ,  $\tilde{C}$  and  $\tilde{F}$  and their F-norm error bounds given by (83)  $\sim$  (86), the  $\mathcal{H}_\infty$ -norm bound of the error system (90) can be explicitly derived from (107) with a given confidence level  $\alpha(n_e^2 + m^2 n_e^2, \chi_\alpha^2)$ .

Next, we will propose an LMI-based approach to deriving an upper model error bound for  $G_e(z)$ . The following Lemma, similar to Lemma 2.1 in [26], provides a useful tool to address the uncertainty of the state-space matrices in innovations model.

*Lemma 2:* Let  $A \in R^{n \times n}$ ,  $Q = Q^T \in R^{n \times n}$ ,  $H_k \in R^{n \times i}$ , and  $E_k \in R^{j \times n}$  ( $k = 1, \dots, K$ ) be given matrices. If there exist  $K$  positive scalars  $\mu_k > 0$  ( $k = 1, \dots, K$ ) and a positive definite matrix  $P > 0$  such that

$$\begin{bmatrix} -P & PA & PH_1 & PH_2 & \cdots & PH_K \\ Q + \sum_{k=1}^K \mu_k E_k^T E_k & 0 & 0 & \cdots & 0 \\ & -\mu_1 I & 0 & \cdots & 0 \\ & & -\mu_2 I & \cdots & 0 \\ & (sym) & & \ddots & \vdots \\ & & & & -\mu_K I \end{bmatrix} < 0 \quad (108)$$

Then the following inequality holds

$$(A + \sum_{k=1}^K H_k F_k E_k)^T P (A + \sum_{k=1}^K H_k F_k E_k) + Q < 0 \quad (109)$$

for each  $F_k$  ( $k = 1, \dots, K$ ) satisfying  $\|F_k\|_2 \leq 1$ .

*Proof:* From (108), applying Schur complement yields

$$\begin{bmatrix} -P + \sum_{k=1}^K \frac{1}{\mu_k} P H_k H_k^T P & PA \\ A^T P & Q + \sum_{k=1}^K \mu_k E_k^T E_k \end{bmatrix} < 0 \quad (110)$$

Multiplying (110) from the left by  $\begin{bmatrix} P^{-1} & 0 \\ 0 & I \end{bmatrix}$  and from the right by  $\begin{bmatrix} P^{-1} & 0 \\ 0 & I \end{bmatrix}$  yields

$$\begin{bmatrix} -P^{-1} + \sum_{k=1}^K \frac{1}{\mu_k} H_k H_k^T & A \\ A^T & Q + \sum_{k=1}^K \mu_k E_k^T E_k \end{bmatrix} < 0 \quad (111)$$

Also, for each  $k$ , we have that

$$\left( \frac{1}{\sqrt{\mu_k}} \begin{bmatrix} H_k \\ 0 \end{bmatrix} - \sqrt{\mu_k} \begin{bmatrix} 0 \\ E_k^T \end{bmatrix} F_k^T \right) \left( \frac{1}{\sqrt{\mu_k}} \begin{bmatrix} H_k \\ 0 \end{bmatrix} - \sqrt{\mu_k} \begin{bmatrix} 0 \\ E_k^T \end{bmatrix} F_k^T \right)^T \geq 0 \quad (112)$$



which can be simplified as

$$\begin{bmatrix} \frac{1}{\mu_k} H_k H_k^T & 0 \\ 0 & \mu_k E_k^T F_k^T F_k E_k \end{bmatrix} \geq \begin{bmatrix} 0 & H_k F_k E_k \\ E_k^T F_k^T H_k^T & 0 \end{bmatrix} \quad (113)$$

The condition  $\|F_k\|_2 \leq 1$ , is equivalent to

$$F_k^T F_k \leq I \quad (114)$$

From (113) and (114), in (111) we have that

$$\begin{bmatrix} -P^{-1} & A + \sum_{k=1}^K H_k F_k E_k \\ A^T + \sum_{k=1}^K E_k^T F_k^T H_k^T & Q \end{bmatrix} < 0 \quad (115)$$

which is equivalent to (109). ■

To compute the  $\mathcal{H}_\infty$  norm of the error system (90), we consider the following minimization problem

$$\begin{aligned} & \min_P \gamma^2 \\ & s.t. \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{F} \end{bmatrix}^T \begin{bmatrix} P & \\ & I \end{bmatrix} \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{F} \end{bmatrix} - \begin{bmatrix} P & \\ & \gamma^2 I \end{bmatrix} < 0 \\ & P > 0 \end{aligned} \quad (116)$$

where

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ 0 & \tilde{A} \end{bmatrix}, \mathcal{B} = \begin{bmatrix} B \\ \tilde{B} \end{bmatrix}, \mathcal{C} = [C \quad -\tilde{C}], \mathcal{F} = [F \quad -\tilde{F}] \quad (117)$$

and

$$\|\delta A\|_F \leq \epsilon_1, \|\delta B\|_F \leq \epsilon_2, \|\delta C\|_F \leq \epsilon_3, \|\delta F\|_F \leq \epsilon_4 \quad (118)$$

This LMI problem is equivalent to  $\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_\infty} \leq \gamma$  for any  $\|\delta A\|_F \leq \epsilon_1, \|\delta B\|_F \leq \epsilon_2, \|\delta C\|_F \leq \epsilon_3, \|\delta F\|_F \leq \epsilon_4$  where  $\epsilon_1 \sim \epsilon_4$  are given by the square roots of the right hand sides of (83)  $\sim$  (86). Then we will use Lemma 2 to address the uncertainty of  $A \sim F$  in (116). For brevity, we denote

$$\tilde{\mathcal{A}} = \begin{bmatrix} \tilde{A} & 0 \\ 0 & \tilde{A} \end{bmatrix}, \tilde{\mathcal{B}} = \begin{bmatrix} \tilde{B} \\ \tilde{B} \end{bmatrix}, \tilde{\mathcal{C}} = [\tilde{C} \quad -\tilde{C}] \quad (119)$$

and

$$\begin{aligned}
H_1 &= \begin{bmatrix} \sqrt{\epsilon_1}I \\ 0 \\ 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} \sqrt{\epsilon_2}I \\ 0 \\ 0 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 0 \\ 0 \\ \sqrt{\epsilon_3}I \end{bmatrix}, \quad H_4 = \begin{bmatrix} 0 \\ 0 \\ \sqrt{\epsilon_4}I \end{bmatrix} \\
F_1 &= \frac{\delta A}{\epsilon_1}, \quad F_2 = \frac{\delta B}{\epsilon_2}, \quad F_3 = \frac{\delta C}{\epsilon_3}, \quad F_4 = \frac{\delta F}{\epsilon_4} \\
E_1 &= [\sqrt{\epsilon_1}I \ 0 \ 0], \quad E_2 = [0 \ 0 \ \sqrt{\epsilon_2}I], \quad E_3 = [\sqrt{\epsilon_3}I \ 0 \ 0], \quad E_4 = [0 \ 0 \ \sqrt{\epsilon_4}I]
\end{aligned} \tag{120}$$

Then the first LMI in (116) becomes

$$\left( \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & 0 \end{bmatrix} + \sum_{k=1}^4 H_k F_k E_k \right)^T \begin{bmatrix} P & \\ & I \end{bmatrix} \left( \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & 0 \end{bmatrix} + \sum_{k=1}^4 H_k F_k E_k \right) - \begin{bmatrix} P & \\ & \gamma^2 I \end{bmatrix} < 0 \tag{121}$$

where it is readily verified that  $\|F_k\|_F \leq 1$  for  $k = 1, 2, 3, 4$ . Due to the fact that  $\|\cdot\|_2 \leq \|\cdot\|_F$ , we also have  $\|F_k\|_2 \leq 1$  for  $k = 1, 2, 3, 4$ . By Lemma 2, we arrive at a suboptimal, convex minimization problem

$$\begin{aligned}
& \min_{P, \mu_k} \bar{\gamma}^2 \\
& s.t. \quad \begin{bmatrix} - \begin{bmatrix} P & \\ & I \end{bmatrix} & \begin{bmatrix} P & \\ & I \end{bmatrix} \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & 0 \end{bmatrix} & \begin{bmatrix} P & \\ & I \end{bmatrix} H_1 & \cdots & \begin{bmatrix} P & \\ & I \end{bmatrix} H_4 \\ & - \begin{bmatrix} P & \\ & \bar{\gamma}^2 I \end{bmatrix} + \sum_{k=1}^4 \mu_k E_k^T E_k & 0 & \cdots & 0 \\ & & -\mu_1 I & \cdots & 0 \\ & & & \ddots & \vdots \\ & & & & -\mu_4 I \end{bmatrix} < 0 \\
& P > 0 \\
& \mu_k > 0 \ (k = 1, 2, 3, 4)
\end{aligned} \tag{122}$$

whose minimum  $\bar{\gamma}^2$  is an upper bound of the  $\gamma^2$  in (116) such that  $\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_\infty} \leq \bar{\gamma}$ .

## V. NUMERICAL RESULTS

In this section, we evaluate our theorems by applying them to two numerical examples in stochastic subspace identification and  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bound estimation, respectively.

### A. Example in Stochastic Subspace Identification

We present Monte Carlo simulation results for the identification of one typical MIMO system (i.e., system with its poles very close to the unit circle) where for most of the identifications, the DARE (9) fails. The results shown afterwards are all from the corrected stochastic subspace identification in which (15) ~ (17) are used to impose the positive realness on the covariance model.

The example corresponds to the following MIMO innovations model:

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 0.874 & 0.8 \\ -0.2 & 0.96 \end{bmatrix} x_k + \begin{bmatrix} 0.18 & 0.85 \\ -0.25 & -0.4 \end{bmatrix} e_k \\ y_k &= \begin{bmatrix} -0.3 & -0.65 \\ 0.76 & -1.1 \end{bmatrix} x_k + e_k \end{aligned} \quad (123)$$

with the covariance  $Q = E[e_k e_k^T] = \begin{bmatrix} 0.075 & 0.037 \\ 0.037 & 0.068 \end{bmatrix}$  and its poles shown in Figure 1.

We identify the simulated system (123) for 200 times in each of which the data size is fixed at  $N = 2500$  and  $m$  in block-Hankel matrix (4) is chosen to be 4. To reflect the performance of the proposed identification procedure, in terms of  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  we define two relative errors  $E_2$  and  $E_\infty$

$$E_2 = \frac{\|G_e(e^{i\omega}) - \tilde{G}_e(e^{i\omega})\|_{\mathcal{H}_2}}{\|G_e(e^{i\omega})\|_{\mathcal{H}_2}}, \quad E_\infty = \frac{\|G_e(e^{i\omega}) - \tilde{G}_e(e^{i\omega})\|_{\mathcal{H}_\infty}}{\|G_e(e^{i\omega})\|_{\mathcal{H}_\infty}} \quad (124)$$

where  $G_e(e^{i\omega})$  is given by (40). Figures 2 shows the two indices  $E_2$  and  $E_\infty$  for 200 identifications. In particular, their respective expectations are 0.5055 and 0.6159, and their respective standard deviations are 0.1936 and 0.2669. Figure 3 shows the frequency response for one identified  $\tilde{G}_e(\omega)$  with  $E_2 = 0.4360$  and  $E_\infty = 0.5272$ . The simulation results show that the improved stochastic subspace identification procedure works well and guarantees a valid model returned even if the system poles are extremely close to unit circle in the case of insufficient data.

### B. Example in the comparison of the asymptotic variance of the transfer function

This example compares the asymptotic variance of the transfer function estimator  $\tilde{G}_e(z) = \tilde{C}(zI - \tilde{A})^{-1}\tilde{B} + \tilde{F}$ , computed using (24), (64) and (67), with the sample variance obtained from

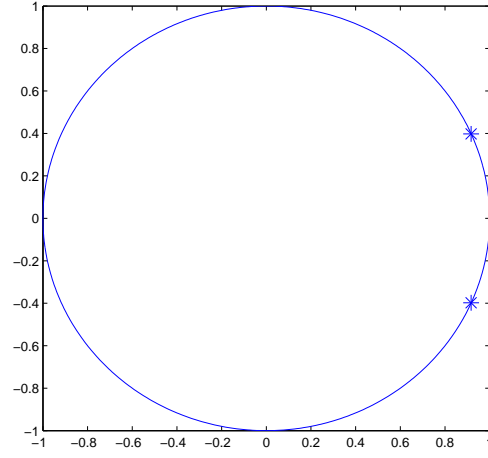


Fig. 1. Poles (\*) of the innovations model (123).

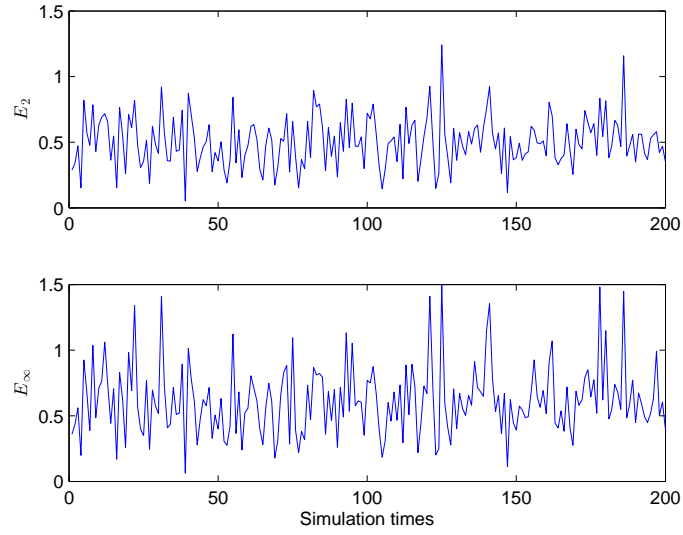


Fig. 2.  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm relative error of the identified system for 200 identifications.

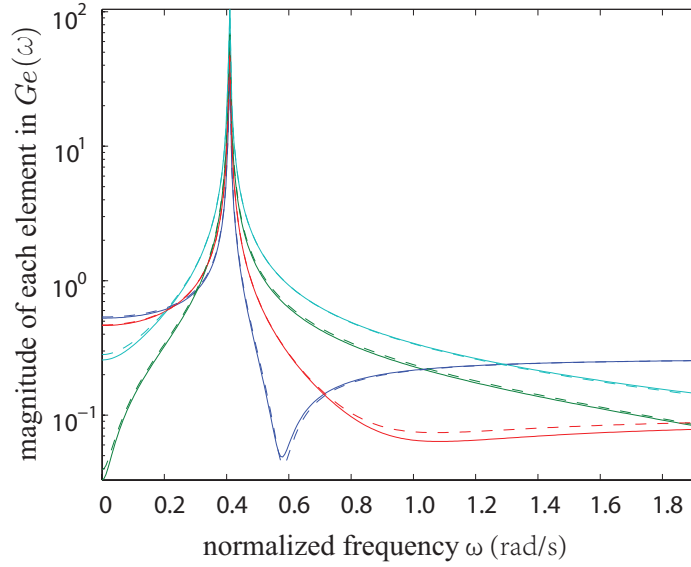


Fig. 3. Frequency response for the original  $G(\omega)$  (solid) and the identified  $\tilde{G}(\omega)$  (dashed).

200 Monte Carlo simulations. Consider the following setup:

$$A = 0.8, \quad C = 0.1, \quad K = 0.35, \quad Q = 0.001$$

, and set the data size to be  $10^4$ .

Figure 4 illustrates that the variance computed from the results of this paper agrees well with the sample variance obtained from the Monte Carlo simulations. Since the second order statistics is used in (24) and only the first-order perturbation is used in (64) and (67), the asymptotic variance requires more data to obtain an accurate estimation than its counterpart in the deterministic case [14].

### C. Example in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Norm Model Error Bound Estimation

In this subsection, we will first identify a stochastic system using the proposed stochastic subspace identification procedure, and then use the estimated state-space matrices to approximately compute the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bounds with a given confidence level. When the system poles are very close to the unit circle, according to our simulations a tight model error bound with a high confidence level requires an enormous data size that conflicts with the practical conditions. Thus, we choose a different simulation example from (123).

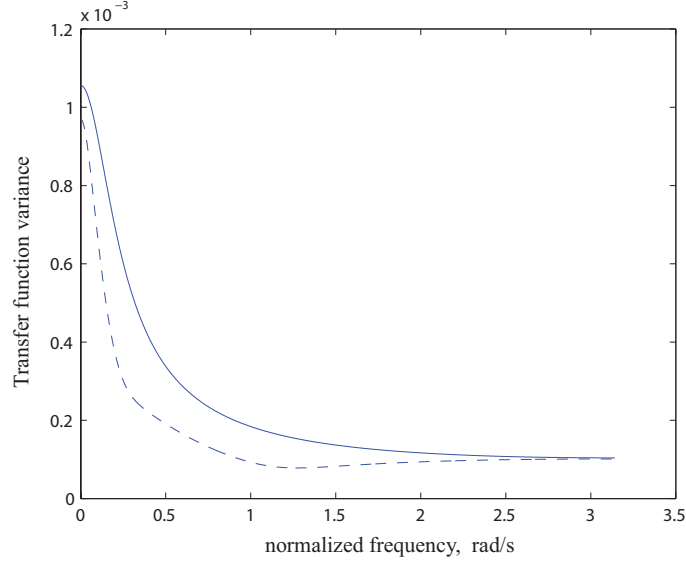


Fig. 4. Asymptotic variance and sample variance (Monte Carlo estimation) versus normalized frequency ( $\omega \in [0, \pi]$ ): the solid line represents asymptotic variance and the dash line represents sample variance.

Consider an MIMO innovations model:

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 0.58 & 0.23 \\ -0.39 & 0.82 \end{bmatrix} x_k + \begin{bmatrix} 0.15 & 0.1 \\ -0.25 & -0.4 \end{bmatrix} e_k \\ y_k &= \begin{bmatrix} -0.3 & -0.65 \\ 0.76 & -1.1 \end{bmatrix} x_k + e_k \end{aligned} \quad (125)$$

with the covariance  $Q = E[e_k e_k^T] = \begin{bmatrix} 0.075 & 0.037 \\ 0.037 & 0.068 \end{bmatrix}$ , the  $\mathcal{H}_2$  norm  $\|G_e(\omega)\|_{\mathcal{H}_2} = 0.5113$ , and  $\mathcal{H}_\infty$  norm  $\|G_e(\omega)\|_{\mathcal{H}_\infty} = 0.9774$ . In the identification procedure, we choose  $m = 4$  in (4) and to get a tight model error bound with a high confidence level, the data size  $N$  is set to be  $1 \times 10^5$ .  $\mathbf{g}$  in (72) is a normally distributed vector, i.e.  $\mathbf{g} \rightarrow \mathcal{N}(0, I_{68})$ . We choose a high confidence level  $\alpha(68, 88.5) = 95.18\%$ , i.e. with the probability 0.9518, such that  $\mathbf{g}^T \mathbf{g} \leq \chi_{0.9518}^2 = 88.5$  in

(72). By Theorem 3, (107) and (122), we compute the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm model error bounds

$$\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_2} \leq 0.1059$$

$$\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_\infty} \leq 0.0848$$

$$\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_\infty} \leq 0.1467$$

respectively, with a confidence level 95.18%, while the true  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm for the error system (90) is

$$\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_2} = 0.0075$$

$$\|G_e(\omega) - \tilde{G}_e(\omega)\|_{\mathcal{H}_\infty} = 0.0164$$

It is shown that (107) is more likely to achieve a tighter  $\mathcal{H}_\infty$  norm model error bound than (122).

## VI. CONCLUSION

In this paper, a new and straightforward LMI-based optimization approach was proposed to impose positive realness on a formerly identified covariance model, guaranteeing a positive definite solution to the DARE (9) and thus a valid innovations model. As can be seen from the numerical results, this approach performs well even if the system poles are very close to unit circle in the case of insufficient data.

Later, we presented a complete asymptotic analysis of the stochastic subspace identification algorithm. It is shown that, if the data size is sufficiently large, the estimated state-space matrices of the covariance model are normally distributed. Also, using Chi-square cumulative distribution function, we derived the asymptotic F-norm error bounds of the estimated state-space matrices in innovations model. By combining these asymptotic results, the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norm error bounds for the identified system are explicitly derived with a given confidence level. Numerical example in  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  model error bound estimation for the identified system is provided to illustrate the performance of the proposed estimation approaches. The tools developed in this paper allow to extract the weighted additive  $\mathcal{H}_\infty$  norm model error bound suited for robust controller design from system identification, which we postpone to future work.

## ACKNOWLEDGMENT

The authors would like to thank...

## REFERENCES

- [1] T. Kailath and A. H. Sayed, *Linear estimation*. NJ: Prentice-Hall: Englewood Cliffs, 2000.
- [2] K. Zhou and J. C. Doyle, *Essentials of robust control*. Upper Saddle River, NJ: Prentice Hall, 1997.
- [3] G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*. Berlin: Springer, 2000.
- [4] P. L. Faurre, *Stochastic realization algorithms*, in: *System identification: Advances and case studies*. Academic Press, New York, NY, 1976.
- [5] P. Van Overschee and B. De Moor, “N4sid: subspace algorithms for the identification of combined deterministic and stochastic systems,” *Automatica*, vol. 30, pp. 75–93, 1994.
- [6] —, *Subspace identification for linear systems: theory-implementation-applications*. Boston, MA: Kluwer, 1996.
- [7] A. Dahlén, A. Lindquist, and J. Mari, “Experimental evidence showing that stochastic subspace identification methods may fail,” *Systems and Control Letters*, vol. 34, pp. 303–312, 1998.
- [8] J. Mari, P. Stoica, and T. McKelvey, “Vector ARMA estimation: a reliable subspace approach,” *IEEE Transaction on Signal Processing*, vol. 48, pp. 1999–2012, 2000.
- [9] I. Goethals, T. Van Gestel, J. Suykens, P. Van Dooren, and B. De Moor, “Identification of positive real models in subspace identification by using regularization,” *IEEE Transaction on Automatic Control*, vol. 48, pp. 1843–1847, 2003.
- [10] R. J. Vaccaro and V. Tomislav, “A solution to the positivity problem in the state-space approach to modeling vector-valued time series,” *Journal of Economic Dynamics and Control*, vol. 17, pp. 401–421, 1993.
- [11] R. Johansson, M. Verhaegen, C. T. Chou, and A. Robertsson, “Residual models and stochastic realization in state-space identification,” *International Journal of Control*, vol. 74, pp. 988–995, 2001.
- [12] L. Ljung, *System identification: theory for the user*. NJ: Prentice-Hall, 1997.
- [13] T. Soderstrom and P. Stoica, *Instrumental variable methods for system identification*. New York: Springer-Verlag, 1983.
- [14] A. Chiuso, “Asymptotic variance of closed-loop subspace identification methods,” *IEEE Transaction on Automatic Control*, vol. 51, pp. 1299–1314, 2006.
- [15] G. C. Goodwin, M. Gevers, and B. M. Ninness, “Optimal model order selection and estimation of model uncertainty for identification with finite data,” in *IEEE Conference on Decision and Control*, Honolulu, HI, 1990.
- [16] B. M. Ninness and G. C. Goodwin, “Robust frequency response estimation accounting for noise and undermodelling,” in *American Control Conference*, Chicago, IL, 1992.
- [17] D. K. De Vries and P. M. Van den Hof, “Quantification of uncertainty in transfer function estimation: a mixed deterministic-probabilistic approach,” in *12th IFAC World Congress*, Sydney, Australia, 1993.
- [18] —, “Quantification of uncertainty in transfer function estimation: a mixed probabilistic-worst case approach,” *Automatica*, vol. 31, no. 4, pp. 543–557, 1995.
- [19] R. G. Hakvoort, “System identification for robust process control. nominal models and error bounds,” Ph.D. dissertation, Technische Univ., Delft, Netherlands, 1994.
- [20] X. Bombois, “Connecting prediction error identification and robust control analysis: a new framework,” Ph.D. dissertation, Université Catholique de Louvain, Belgique, 2000.
- [21] P. H. Janssen and P. Stoica, “On the expectation of the product of four matrix-valued gaussian random variables,” *IEEE Transaction on Automatic Control*, vol. 33, pp. 867–870, 1988.
- [22] P. J. Brockwell and R. A. Davis, *Time series, theory and methods*. New York: Springer-Verlag, 1991.
- [23] J. P. Delmas and Y. Meurisse, “Asymptotic performance analysis of DOA finding algorithms with temporally correlated narrowband signals,” *IEEE Transaction on Signal Processing*, vol. 48, pp. 2669–2674, 2000.



- [24] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2820–2830, 2002.
- [25] M. M. Konstantinov, P. H. Petkov, and N. D. Christov, "Perturbation analysis of the discrete Riccati equation," *Kybernetika*, vol. 29, pp. 18–29, 1993.
- [26] L. Xie and Y. C. Soh, "Control of uncertain discrete-time systems with guaranteed cost," in *IEEE Conference on Decision and Control*, 1993, pp. 56–61.



**Quan Li** Biography text here.



**Jeff Scruggs** Biography text here.